NZ00/176

Intellectual Property Office of New Zealand

4

PRIORITY 10/069527
DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

CERTIFICATE

This certificate is issued in support of an application for Patent registration in a country outside New Zealand pursuant to the Patents Act 1953 and the Regulations thereunder.

I hereby certify that annexed is a true copy of the Provisional Specification as filed on 7 September 1999 with an application for Letters Patent number 337688 made by THE HORTICULTURE AND FOOD RESEARCH INSTITUTE OF NEW ZEALAND LIMITED.

Dated 19 September 2000.

Neville Harris
Commissioner of Patents

337688

Patents Form No. 4

PATENTS ACT 1953

PROVISIONAL SPECIFICATION

SEEDLESS FRUIT PRODUCTION

We, **THE HORTICULTURE AND FOOD RESEARCH INSTITUTE OF NEW ZEALAND LIMITED**, a New Zealand company of Batchelar Research Centre, Highway 57, Palmerston North, New Zealand do hereby declare this invention to be described in the following statement:

SEEDLESS FRUIT PRODUCTION

FIELD OF THE INVENTION

The invention provides plants that produce seedless or sterile fruit.

BACKGROUND TO THE INVENTION

The production of seedless or parthenocarpic fruit is a desirable trait for commercially grown cultivars. Seedless fruit are more convenient than seeded fruit to consumers. Furthermore parthenocarpic fruit trees can be cropped without pollination, which reduces dependence on bees, pollinator varieties and warm weather at flowering. The absence of pollen is also advantageous so as to alleviate environmental concerns regarding the transfer of transgenes to non-transgenics by cross-pollination.

Seedless fruit cultivars can also avoid or reduce biennial bearing tendencies that have been attributed to the inhibition of flower bud formation by developing seeds in apple (Chan and Cain, 1967). Seedless apple fruit is also much less susceptible to codling moth, a major pest on apple trees, compared to seeded fruit (Goonewardene *et al.*, 1984).

The applicants have now identified and isolated a reproductive gene which encodes a peptide involved in the reproductive (seed-producing) cycle of fruiting plants, particularly apple trees. It is broadly towards this gene, to its homologs in other fruiting plants and to the modulation of its expression/function within fruiting plants that the present invention is directed.

SUMMARY OF THE INVENTION

In a first aspect, the present invention broadly provides a fruiting plant which contains a polynucleotide encoding a peptide having the *MdPI* amino acid sequence of Figure 2 or a functionally equivalent variant thereof, and which has been genetically modified to produce seedless or sterile fruit.

As used herein, "fruiting plant" means a plant in which the fruit is formed from the ovary and the fused bases of sepals, petals and stamen.

Conveniently, the plant is one which produces a pome fruit.

Preferably, the genetic modification reduces the functional expression of said peptide (ie. the amount of the peptide which is expressed and functional within the plant).

In a second aspect the invention provides a genetically transformed fruiting plant in which expression of the polynucleotide transcript corresponding to nucleotides 1 to 842 of the sequence of Figure 2, or its functional equivalent, has been disrupted.

Disruption of gene expression may be by mutation (such as frameshift, deletion, insertion or knockout mutations) of the gene itself or of its regulatory elements, down-regulation (such as antisense, co-suppression) or any other method known to those skilled in the art by which aberrant or reduced expression of the gene may be achieved (e.g. Montgomery and Fire, 1998).

Disruption may be specifically caused by down-regulation of expression of *MdPI*, *MdAP3*, or both.

The invention includes seedless or sterile fruit produced by a plant as defined above.

In a further embodiment, the invention provides a polynucleotide which encodes a peptide having the *MdPI* amino acid sequence of Figure 2 or a variant thereof.

Most preferably, said polynucleotide includes part or all of the nucleotide sequence of Figure 2.

Preferably, the polynucleotide is DNA.

The invention further provides a DNA construct which includes a polynucleotide as defined above.

More particularly, the invention provides a DNA construct comprising, in the 5'-3' direction:

(a) a promoter sequence;

(b) an open reading frame polynucleotide coding for the peptide having the *MdPI* amino acid sequence of Figure 2 or a functionally equivalent variant thereof; and

(c) a termination sequence.

In one embodiment, the open reading frame is in a sense orientation.

In an alternative embodiment, the open reading frame is in an anti-sense orientation.

In still a further embodiment, the invention provides a DNA construct comprising, in the 5'-3' direction:

(a) a promoter sequence;

(b) a non-coding region of a gene coding for the peptide having the *MdPI* amino acid sequence of Figure 2 or a functionally equivalent variant thereof; and

(c) a termination sequence.

Once again, the non-coding region can be in a sense or anti-sense orientation.

In yet a further embodiment, the invention provides a DNA construct comprising, in the 5'-3' direction:

(a) a promoter sequence;

(b) a polynucleotide comprising a nucleotide sequence complementary to at least part of a sequence coding for the peptide having the *MdPI* amino acid sequence of Figure 2 or a functionally equivalent variant thereof; and

(c) a termination sequence.

Preferably, in each embodiment, the construct further includes a marker for identification of transformed cells.

Similar constructs can also be provided including a polynucleotide which encodes part or all of *MdAP3* having the sequence of Figure 6.

In still a further aspect, the invention provides a transgenic fruiting plant cell which includes a DNA construct as defined above.

In a further aspect the invention provides a method of hybridisation employing a polynucleotide construct comprising nucleotides 1 to 842 of the sequence of Figure 2 or part thereof or comprising nucleotides 1 to 982 of the sequence of Figure 6 or part thereof, to isolate or detect the level of expression of the encoding gene or its intra- or interspecific homologs in other fruiting plants.

DESCRIPTION OF THE DRAWINGS

While the invention is broadly defined as above, those persons skilled in the art will appreciate that it is not limited thereto and that it also includes embodiments of which the following description provides examples. In addition, the present invention will be better understood from reference to the accompanying drawings in which:

Figure 1 shows the phenotype of wild type and Rae Ime apple flowers and fruit.

(a) normal apple flowers showing sepals, petals, stamens and styles.

(b) a normal 5-week-old apple fruit showing five carpels with 0 to 2 seeds per carpel.

(c) Rae Ime flowers with no petals or stamens but with increased numbers of styles.

(d) cross sections at the lower part (left) and upper part of a 5-week-old Rea Ime fruit, showing two whorls of carpels without seed.

(e) top of Rea Ime fruit showing two whorls of calyxes.

(f) top of normal apple fruit showing a whorl of calyxes.

(g) mature fruit of Rea Ime with size of 5 cm wide and no seed.

Figure 2 shows the sequence of *MdPI*. The cDNA sequences and deduced amino acid sequences of *MdPI* isolated from Granny Smith apple and *Arabidopsis PI* protein are shown in the up, middle and low lines respectively. Gene specific PCR primers are underlined. Primer directions are indicated with horizontal arrows. Intron positions are indicated with vertical arrows.

Figure 3 shows a Northern blot analysis of apple RNA sample using *MdPI* cDNA as a probe. RNA sample were prepare from ovaries (1), sepals (2), young leaves (3), skin (4), cortex (5) and core (6) tissue of 4-week-old fruit of Granny Smith, 1-week-old fruit (7), flower peduncles (8), stamens (9), petals (10) of Granny Smith (12), flower buds of Rea Ime (11), and flower buds of Granny Smith (12).

Figure 4 shows a Southern analysis of apple genomic DNA using *MdPI* cDNA as a probe. DNA of Rea Ime (Ri) and Granny Smith (Gs) were digested with EcoRI (E) and HindIII (H).

Figure 5 shows the identification of an insertion in *MdPI* of Rea Ime.

(a) Genomic DNA fragments were amplified using primers P3 and P7 from Rea Ime (Ri) and Granny Smith (Gs).

(b) Southern blot made from the gel shown in (a) was probed with the cDNA of *MdPI*.

(c) An alignment of part of genomic sequences from the 5' end of *MdPI* from Rea Ime (Ri 5'), *MdPI* from Granny Smith (Gs) and 3' end of MdPI from Rea Ime (Ri 3'). The insertion site was identified as shown by the arrow.

Figure 6 shows the cDNA and deduced amino acid sequences of *MdAP3*.

DESCRIPTION OF THE INVENTION

As broadly outlined above, the applicants have identified a peptide which is involved in fruiting plant reproduction, together with the gene coding therefor. The specific peptide and gene are from a plant which produces pome fruit, *Malus x domestica*.

The amino acid sequence of one peptide, *MdPI*, and its encoding nucleotide sequence are given in Figure 2. It will however be appreciated that the invention is not restricted only to the peptide/polynucleotide having the specific amino acid/nucleotide sequence given in Figure 2. Instead, the invention also extends to functionally equivalent variants of the peptide/polynucleotide of Figure 2.

The term "polynucleotide(s)" as used herein means a single or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases and includes DNA and corresponding RNA molecules, including HnRNA and mRNA molecules, both sense and anti-sense strands, and comprehends cDNA, genomic DNA and recombinant DNA, as well as wholly or partially synthesized polynucleotides. An HnRNA molecule contains introns and corresponds to a DNA molecule in a generally one-to-one manner. An mRNA molecule corresponds to an HnRNA and DNA molecule from which the introns have been excised. A polynucleotide may consist of an entire gene, or any portion thereof. Operable anti-sense polynucleotides may comprise a fragment of the corresponding polynucleotide, and the definition of "polynucleotide" therefore includes all such operable anti-sense fragments.

The phrase "functionally equivalent variants" recognises that it is possible to vary the amino acid/nucleotide sequence of a peptide while retaining substantially equivalent functionality. For example, a peptide can be considered a functional equivalent of another peptide for a specific function if the equivalent peptide is immunologically cross-reactive with and has at least substantially the same function as the original peptide. The equivalent can be, for example, a fragment of the peptide, a fusion of the peptide with another peptide or carrier, or a fusion of a fragment which additional amino acids. For example, it is possible to substitute amino acids in a sequence with equivalent amino acids using conventional techniques. Groups of amino acids normally held to be equivalent are:

(a) Ala, Ser, Thr, Pro, Gly;

(b) Asn, Asp, Glu, Gln;

(c) His, Arg, Lys;

(d) Met, Leu, Ile, Val; and

(e) Phe, Tyr, Trp.

Equally, nucleotide sequences encoding a particular product can vary significantly simply due to the degeneracy of the nucleic acid code.

Variants can have a greater or lesser degree of homology as between the variant amino acid/nucleotide sequence and the original.

Polynucleotide or polypeptide sequences may be aligned, and percentage of identical nucleotides in a specified region may be determined against another sequence, using computer algorithms that are publicly available. Two exemplary algorithms for aligning and identifying the similarity of polynucleotide sequences are the BLASTN and FASTA algorithms. The similarity of polypeptide sequences may be examined using the BLASTP algorithm. Both the BLASTN and BLASTP software are available on the NCBI anonymous FTP server (ftp://ncbi.nlm.nih.gov) under /blast/executables/. The BLASTN algorithm version 2.0.4 [Feb-24-1998], set to the default parameters described in the documentation of variants according to the present invention. The use of the BLAST family of algorithms, including BLASTN and BLASTP, is described at NCBI's website at URL http://www.ncbi.nlm.nih.gov/BLAST/newblast.html and in the publication of Altschul, Stephen F., *et al.* (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res.* 25:3389-34023. The computer algorithm FASTA is available on the Internet at the ftp site ftp://ftp.virginia.edu/pub/fasta/. Version 2.0u4, February 1996, set to the default parameters described in the documentation and distributed with the algorithm, is also preferred for use in the determination of variants according to the present invention. The use of the FASTA algorithm is described in W. R. Pearson and D. J. Lipman, "Improved Tools for Biological Sequence Analysis", *Proc. Natl. Acad. Sci. USA* 85:2444-2448 (1988) and W. R. Pearson, "Rapid and Sensitive Sequence Comparison with FASTP and FASTA, "*Methods in Enzymology* 183:63-98 (1990).

The following running parameters are preferred for determination of alignments and similarities using BLASTN that contribute to E values (as discussed below) and percentage identity: Unix running command: blastall -p blastn -d embldb -e 10 -G 1 -E 1 -r 2 -v 50 -b 50 -I queryseq -o results; and parameter default values:

-p Program Name [String]

-d Database [String]

-e Expectation value (E) [Real]

-G Cost to open a gap (zero invokes default behaviour) [Integer]

-E Cost to extend a cap (zero invokes default behaviour) [Integer]

-r Reward for a nucleotide match (blastn only) [Integer]

-v Number of one-line descriptions (V) [Integer]

-b Number of alignments to show (B) [Integer]

-i Query File [File In]

-o BLAST report Output File [File Out] Optional

For BLASTP the following running parameters are preferred: blastall -p blastp -d swissprotdb -e 10 -G 1 -E 1 -v 50 -b 50 -I queryseq -o results

-p Program Name [String]

-d Database [String]

-e Expectation value (E) [Real]

-G Cost to open a gap (zero invokes default behaviour) [Integer]

-E Cost to extend a cap (zero invokes default behaviour) [Integer]

-v Number of one-line descriptions (v) [Integer]

-b Number of alignments to show (b) [Integer]

-i Query File [File In]

-o BLAST report Output File [File Out] Optional

The "hits" to one or more database sequences by a queried sequence produced by BLASTN, BLASTP, FASTA, or a similar algorithm, align and identify similar portions of sequences. The hits are arranged in order of the degree of similarity and the length of sequence overlap. Hits to a database sequence generally represent an overlap over only a fraction of the sequence length of the queried sequence.

The BLASTN and FASTA algorithms also produce "Expect" or E values for alignments. The E value indicates the number of hits one can "expect" to see over a certain number of contiguous sequences by chance when searching a database of a certain size. The Expect value is used as a significance threshold for determining whether the hit to a database, such as the preferred EMBL database, indicates true similarity. For example, an E value of 0.1 assigned to a hit is interpreted as meaning that in a database of the size of the EMBL database, one might expect to see 0.1 matches over the aligned portion of the sequence with a similar score simply by chance. By this criterion, the aligned and matched portions of the sequences

then have a 90% probability of being the same. For sequences having an E value of 0.01 or less over aligned and matched portions, the probability of finding a match by chance in the EMBL database is 1% or less using the BLASTN or FASTA algorithm.

According to one embodiment, "variant" polynucleotides, with reference to each of the polynucleotides of the present invention, preferably comprise sequences having the same number or fewer nucleic acids than each of the polynucleotides of the present invention and producing an E value of 0.01 or less when compared to the polynucleotide of the present invention. That is, a variant polynucleotide is any sequence that has at least a 99% probability of being the same as the polynucleotide of the present invention, measured as having an E value of 0.01 or less using the BLASTN or FASTA algorithms set at the parameters discussed above.

Variant polynucleotide sequences will generally hybridize to the recited polynucleotide sequence under stringent conditions. As used herein, "stringent conditions" refers to prewashing in a solution of 6X SSC, 0.2% SDS; hybridizing at 65°C, 6X SSC, 0.2% SDS overnight; followed by two washes of 30 minutes each in 1X SSC, 0.1% SDS at 65°C and two washes of 30 minutes each in 0.2X SSC, 0.1% SDS at 65°C.

It is of course expressly contemplated that homologs to *MdPI* exist in other fruiting plants. Such homologs are also "functionally equivalent variants" of *MdPI* as the phrase is used herein.

DNA sequences from fruiting plants other than *Malus x domestica* which are homologs of *MdPI* may be isolated by high throughput sequencing of cDNA libraries prepared from such plants. Alternatively, oligonucleotide probes based on the sequences for *MdPI* provided in Figure 2 can be synthesized and used to identify positive clones in either cDNA or genomic DNA libraries from other plants by means of hybridization or PCR techniques. Probes should be at least about 10, preferably at least about 15 and most preferably at least about 20 nucleotides in length. Hybridization and PCR techniques suitable for use with such oligonucleotide probes are well known in the art. Positive clones may be analyzed by restriction enzyme digestion, DNA sequencing or the like.

The polynucleotides of the present invention may be generated by synthetic means using techniques well known in the art. Equipment for automated synthesis of oligonucleotides is commercially available from suppliers such as Perkin Elmer/Applied Biosystems Division (Foster City, CA) and may be operated according to the manufacturer's instructions.

The primary importance of identification of the peptide/polynucleotides of the invention is that they enable the reproductive (seed-producing) capacity of fruiting plants to be modulated. This modulation will generally involve a reduction in the functional expression (silencing) of the reproductive peptide.

Any conventional technique for effecting this can be employed. Intervention can occur post-transcriptionally or pre-transcriptionally. Further, intervention can be focused upon the gene itself or on regulatory elements associated with the gene and which have an effect on expression of the encoded peptide. "Regulatory elements" is used here in the widest possible sense and includes other genes which interact with the gene of interest. For example, intervention which targets expression of *MdAP3* peptide is contemplated. *MdAP3* is functionally related to *MdPI* such that down-regulation of *MdAP3* expression will in turn down-regulate *MdPI* (see Jack *et al* (1992) and Goto & Meyerowitz (1994)).

The cDNA and deduced amino acid sequences for *MdAP3* are shown in Figure 6.

Pre-transcription intervention can involve mutation of the gene itself or of its regulatory elements. Such mutations can be point mutations, frameshift mutations, insertion mutations or deletion mutations.

Examples of post-transcription interventions include co-suppression or anti-sense strategies, a dominant negative approach, or techniques which involve ribozymes to digest, or otherwise be lethal to, RNA post-transcription of the target gene.

Co-suppression can be effected in a manner similar to that discussed, for example, by Napoli *et al* (Plant Cell 2:279-290, 1990) and de Carvalho Niebel *et al* (Plant Cell 7:347-258, 1995). In some cases, it can involve overexpression of the gene of interest through use of a constitutive promoter. It can also involve transformation

of a plant with a non-coding region of the gene, such as an intron from the gene or 5'-non-coding leader sequences.

Anti-sense strategies involve expression or transcription of DNA with the expression/transcription product being capable of interfering with translation of mRNA transcribed from the target gene. This will normally be through the expression/transcription product hybridising to and forming a duplex with the target mRNA.

The expression/transcription product can be a relatively small molecule and still be capable of disrupting mRNA translation. However, the same result is achieved by expressing the target gene in an anti-sense orientation such that the RNA produced by transcription of the anti-sense oriented gene is complementary to all or part of the endogenous target mRNA.

Anti-sense strategies are described generally by Robinson-Benion *et al.*, (1995), Anti-sense techniques, *Methods in Enzymol.* 254(23):363-375 and Kawasaki *et al.*, (1996), *Artific. Organs* 20 (8): 836-848.

Dominant negative approaches involve the expression of a modified DNA binding/activating protein which includes a DNA binding domain but not a activator domain. The result is that the protein binds to DNA as intended but fails to activate, while at the same time blocking the binding of the DNA binding/activating peptides which normally bind to the same site.

The ribozyme approach to regulation of peptide expression involves inserting appropriate sequences or subsequences (eg. DNA or RNA) in ribozyme constructs (McIntyre CL, Manners JM, *Transgenic Res.* 5(4):257-262, 1996). Ribozymes are synthetic RNA molecules that comprise a hybridizing region complementary to two regions, each of which comprises at least 5 contiguous nucleotides of a mRNA molecule encoded by one of the inventive polynucleotides. Ribozymes possess highly specific endonuclease activity, which autocatalytically cleaves the mRNA.

To give effect to the above strategies, the invention also provides DNA constructs. The constructs include the intended DNA (such as the gene of the invention in anti-

sense orientation or a polynucleotide encoding the appropriate DNA binding domain or ribozyme), a promoter sequence and a termination sequence, operably linked to the DNA sequence to be transcribed, which control expression of the gene. The promoter sequence is generally positioned at the 5' end of the DNA sequence to be transcribed, and is employed to initiate transcription of the DNA sequence. Promoter sequences are generally found in the 5' non-coding region of a gene but they may exist in introns (Luehrsen, K.R., *Mol. Gen. Genet.* 225:81-93, 1991) or in the coding region. When the construct includes an open reading frame in a sense orientation (for co-suppression through over-expression) the promoter sequence also initiates translation of the open reading frame. For DNA constructs comprising either an open reading frame in an anti-sense orientation or a non-coding region, the promoter sequence generally consists only of a transcription initiation site having a RNA polymerase binding site.

A variety of promoter sequences which may be usefully employed in the DNA constructs of the present invention are well known in the art. The promoter sequence, and also the termination sequence, may be endogenous to the target *Malus* plant host or may be exogenous, provided the promoter is functional in the target host. For example, the promoter and termination sequences may be from other plant species, plant viruses, bacterial plasmids and the like. Preferably, promoter and termination sequences are those endogenously associated with the reproductive genes.

Factors influencing the choice of promoter include the desired tissue specificity of the construct, and the timing of transcription and translation. For example, constitutive promoters, such as the 35S Cauliflower Mosaic Virus (CaMV 35S) promoter, will affect the activity in all parts of the plant. Use of a tissue specific promoter will result in production of the desired sense or antisense RNA only in the tissue of interest. With DNA constructs employing inducible promoter sequences, the rate of RNA polymerase binding and initiation can be modulated by external stimuli, such as light, heat, anaerobic stress, alteration in nutrient conditions and the like. Temporally regulated promoters can be employed to effect modulation of the rate of RNA polymerase binding and initiation at a specific time during development of a transformed cell. Preferably, the original promoters from the gene in question, or promoters from a specific tissue-targeted gene in the organism to be

transformed are used. Other examples of promoters which may be usefully employed in the present invention include, mannopine synthase (mas), octopine synthase (ocs) and those reviewed by Chua *et al.* (Science, 244:174-181, 1989).

The termination sequence, which is located 3' to the DNA sequence to be transcribed, may come from the same gene as the promoter sequence or may be from a different gene. Many termination sequences known in the art may be usefully employed in the present invention, such as the 3' end of the *Agrobacterium tumefaciens* nopaline synthase gene. However, preferred termination sequences are those from the original gene or from the target *Malus* species to be transformed.

The DNA constructs of the present invention may also contain a selection marker that is effective in plant cells, to allow for the detection of transformed cells containing the construct. Such markers, which are well known in the art, typically confer resistance to one or more toxins. One example of such a marker is the NPTII gene whose expression results in resistance to kanamycin or hygromycin, antibiotics which is usually toxic to plant cells at a moderate concentration (Rogers *et al.*, in *Methods for Plant Molecular Biology*, A Weissbach and H Weissbach eds, Academic Press Inc., San Diego, CA (1988)). Alternatively, the presence of the desired construct in transformed cells can be determined by means of other techniques well known in the art, such as Southern and Western blots.

Techniques for operatively linking the components of the inventive DNA constructs are well known in the art and include the use of synthetic linkers containing one or more restriction endonuclease sites as described, for example, by Maniatis *et al.*, (*Molecular Cloning: A Laboratory Manual*, Cold Spring Harbour Laboratories, Cold Spring Harbour, NY, 1989). The DNA construct may be linked to a vector having at least one replication system, for example, *E. coli*, whereby after each manipulation, the resulting construct can be cloned and sequenced and the correctness of the manipulation determined.

The DNA constructs of the present invention may be used to transform a variety of fruiting plants. In a preferred embodiment, the DNA constructs are employed to transform apple and its related species such as pear.

As discussed above, transformation of a fruiting plant with a DNA construct including an open reading frame coding for a peptide encoded by a DNA sequence of the invention wherein the open reading frame is orientated in a sense direction can, in some cases, lead to a decrease in expression of the peptide by co-suppression. Transformation of the plant with a DNA construct comprising an open reading frame in an anti-sense orientation or a non-coding (untranslated) region of a gene will lead to a decrease in the expression of the peptide in the transformed plant.

Techniques for stably incorporating DNA constructs into the genome of target fruiting plants are well known in the art and include *Agrobacterium tumefaciens* mediated introduction, electroporation, protoplast fusion, injection into reproductive organs, injection into immature embryos, high velocity projectile introduction and the like. The choice of technique will depend upon the target plant to be transformed.

Once the cells are transformed, cells having the DNA construct incorporated into their genome may be selected by means of a marker, such as the kanamycin resistance marker discussed above. Transgenic cells may then be cultured in an appropriate medium to regenerate whole plants, using techniques well known in the art. In the case of protoplasts, the cell wall is allowed to reform under appropriate osmotic conditions. In the case of seeds or embryos, an appropriate germination or callus initation medium is employed. For explants, an appropriate regeneration medium is used.

For a review of regeneration of trees, see Dunstan *et al.*, Somatic embryogenesis in woody plants. In: Thorpe, T.A. ed. 1995: *in vitro* embryogenesis of plants. Vol 20 in Current Plant Science and Biotechnology in Agriculture, Chapter 12, pp. 471-540.

The resulting transformed fruiting plants may be reproduced sexually or asexually, using methods well known in the art, to give successive generations of transgenic plants.

The nucleotide sequence information provided herein will also be useful in programs for identifying nucleic acid variants from fruiting plants and for pre-selecting plants with mutations in *MdPI, MdAP3* or their equivalents which renders those plants

useful in an accelerated breeding program to produce seedless fruit. More particularly, the nucleotide sequence information provided herein may be used to design probes and primers for probing or amplification of *MdPI*, *MdAP3* or variants thereof. An oligonucleotide for use in probing or PCR may be about 30 or fewer nucleotides in length. Generally, specific primers are upwards of 14 nucleotides in length. For optimum specificity and cost effectiveness, primers or 16-24 nucleotides in length are preferred. Those skilled in the art are well versed in the design of primers for use in processes such as PCR.

If required, probing can be done with entire restriction fragments of the gene disclosed herein. Naturally, sequences based upon Figure 2, or Figure 6 or the complements thereof can be used.

Such probes and primers also form aspects of the present invention.

Probing may employ the standard Southern blotting technique. For instance, DNA may be extracted from cells and digested with different restriction enzymes. Restriction fragments may then be separated by electrophoresis on an agarose gel, before denaturation and transfer to a nitrocellulose filter. Labelled probes may be hybridised to the DNA fragments on the filter and binding determined. DNA for probing may be prepared from RNA preparations from cells. Probing may optionally be done by means of so-called "nucleic acid chips" (see Marshall and Hodgson (1998)).

The invention will now be illustrated with reference to the following non-limiting experiments.

EXPERIMENTAL

Methods and Materials

Cloning MdPI using PCR approaches

Total RNA was isolated from 'Granny Smith' apple flowers using the method described by Chang *et al* (1993). Poly(A) mRNA was purified from the total RNA using the mRNA Purification Kit (Pharmacia, Sweden). cDNA was synthesized from

the mRNA using the ZAP cDNA Synthesis Kit (Stratagene, CA, USA). DNA fragments were amplified from templates of cDNA using two degenerative PCR primers P1 CGGAATTCATGGGNMGNGGNAARRT-3' and P2 CGCTCGAGGATCCGGYTGNATNGGYTGNAC-3' (N=ATGC, M=AC, R=AG, Y=CT). The primers were designed according the conserved amino acid sequences MGRGKI in the MADS-box domain and VQPM/IQP in the C-terminal region (Fig. 2) in an alignment of PI, GLOBSA, FBP3, SLM2 and pMADS2. The underlined Eco RI and Bam HI sites were used for cloning the PCR products. The PCR amplification conditions were as follows: initial denaturation at 94°C for 4 min; then 35 cycles of 94°C for 1 min, 60°C for 1 min, and 72°C for 3 min, and with a final extension of 5 min at 72°C. Several bands were detected from the PCR on agarose gels and DNA in a band of the expected size (630 bp) was cloned into Bluescript SK (Stratagene, CA, USA) following Eco RI and Bam HI digestion. After the sequences of cloned fragments were determined, two nested PCR primers, P3 and P4 (Fig. 2) were designed using the sequences within the K-box and were used to amplify the 3' region of MdPI cDNA together with a 3' RACE primer GAGAGAGAACTAGTCTCGAG-3'. The PCR conditions were the same as above except for the anneal temperature reduced to 50°C. The amplified fragments were cloned into pGEM-T EASY Vector (Promega).

Genomic fragments of MdPI were amplified using primers P5 and P6, P3 and P7 (Fig. 2). PCR conditions were: initial denaturation at 94°C for 2 min; then 10 cycles of 94°C for 15 sec, 58°C for 30 sec; and 20 cycles of 94°C for 15 sec, 58°C for 30 sec and 68°C for 5 min plus cycle elongation of 20 sec for each cycle; and with a final extension of 5 min at 86°C. The amplified fragments were cloned into pGEM-T EASY Vector. Expand High Fidelity PCR System (Boehringer Mannheim) was used for all PCR experiments.

DNA sequence determination

Nucleotide sequences of MdPI clones were determined using the automatic sequencer ABI PRISM model 377(CA, USA) with universal forward and reverse primers. To obtain complete sequences, gene specific primers were designed and ordered from BRL Life Technologies.

Northern and Southern analysis using MdPI on apple tissues

Total RNA was isolated as described by Chang *et al* (1993) from 'Granny Smith' and Rae Ime apple tissues. Northern blots were prepared as described by Dong *et al* (1997). The northern blot contained RNA isolated from expanding leaves, unopened flowers, and fruit at 2 days and 1, 4 and 8 weeks following hand-pollination. At 4 weeks after pollination, apple fruit is large enough to allow for easy separation into the three main tissue types namely; core, cortex and skin.

DNA was isolated from leaf tissue of Granny Smith and Rae Ime using the method of Rogers and Bendich (1988). Southern blots were prepared by digesting apple DNA (approximately 20 µg per lane) with EcoRI or HindIII, separating DNA fragments on 0.7% agarose gel and transferring them to Hybond-N+ membrane.

Northern and Southern blots were probed with 32P-dCTP labelled PI cDNA clone lacking the MADS-box sequence to significantly reduce cross hybridization32P-dCTP labelled MADS-box DNA fragments. The blots were hybridized in 0.5M NaPO4 buffer (pH 7.2) with 1 mM EDTA and 7% SDS at 65°C and washed using 0.4x SSC and 0.2% SDS at 65°C. Hybridisation signals were detected using a Storm 840 PhosphorImager (Molecular Dynamics, Sunnyvale, California, USA).

Results/Discussion

Flowers of the majority of apple taxa bear 5 sepals, 5 petals, 9-20 stamens (Fig. 1a) and an inferior ovary. These flowers develop into a pome fruit that consists of fleshy cortex tissue derived from the fused bases of sepals, petals and stamens, and the core tissue derived from fertilised ovary containing 5 carpels and up to 10 seeds (Pratt, 1988) (Fig. 1b). In contrast, flowers of Rae Ime show no petal or stamens but increased numbers of styles (Fig. 1c). These flowers develop into seedless fruit without the need for pollination. These seedless fruit have two whorls of carpels, five carpels in the lower whorl and 9 to 10 in the upper whorl (Fig. 1d). The fruit also has duplicated whorls of calyxes (Fig. 1e) that are the remains of sepals, compared to one calyx whorl in a normal apple (Fig. 1f). The mature seedless fruit are close to normal apple fruit size, but the fruit cores are relatively smaller (Fig.1g).

Several apple varieties, such as Spencer Seedless and Wellington Bloomless (Tobutt, 1994), have been described with a very similar flower and fruit structure to that of Rae Ime. Anatomy studies of the vascular connections show that the upper whorl of carpels has been transformed from the stamens and the second whorl of sepals from petals (Brase, 1937). In the *Arabidopsis pi* and *ap3* mutants, flowers have no petals or stamens but have double the number of sepals and carpels (Goto and Meyerowitz, 1994; Jack *et al.*, 1992).

A difference between Rae Ime apple and *pi Arabidopsis* is that the former produces parthenocarpic fruit but the latter does not. Up to 6 apple varieties have been recorded to produce apetalous flowers and parthenocarpic fruit in different countries. Many of these records can be traced back to several centuries ago (Brase, 1937; Tobutt, 1994). This indicates some of the apple mutants may have occurred independently.

Genetic analysis has been performed using two apetalous/parthenocarpic varieties, Spencer Seedless and Wellington Bloomless. Crossing pollen from the cultivar Wijcik with normal flowers to Wellington Bloomless generates hybrids that all produce normal flowers. Crossing the pollen from these hybrids to Spencer Seedless generates plants of which half produce normal flowers and half produce apetalous flowers and parthenocarpic fruit (Tobutt, 1994). This result indicates that a single recessive gene controls apetalous flower development and subsequently parthenocarpic fruit formation. This result also indicates that mutations in Spencer Seedless and Wellington Bloomless are different alleles at the same locus. Independently isolated mutant alleles at the same locus are good evidences for a single gene being involved in the development of apetalous flower and parthenocarpic fruit in these apple mutants.

DNA fragments of 630bp have been amplified from apple flower cDNA using degenerative PCR primers against conserved sequences in the MADS-box and in the C-terminal region of *PI* and its homologues. After these DNA fragments were cloned, 6 random clones were sequenced and found to contain the same sequences. The cloned cDNA sequences started from the first presumed ATG start coden, contained MADS-box, K-box and most of the C-terminal region and had high homology to *PI*. The C-terminal and the 3' un-translated regions were further amplified using two

nested PCR primers within the K-box and a 3' RACE primer. Six clones containing the 3' fragments were sequenced and found to contain the same sequences overlapping with those in the 5' clone. Sequences from the 5' and 3' clone were assembled together and shown in Fig. 2. These sequences show highest homology to *PI* and its homologues (GLOBOSA, FBP3, SLM2 and pMADS2) in Blast searches carried out in GeneBank. The putative apple *PI* homologue was named as *MdPI* having a deduced amino acid sequence identity of 64% to that of *Arabidopsis* PI protein.

MdPI is found to be highly expressed in petals and stamens as determined through northern analysis. Expression in other apple tissues, including sepals and ovaries, is either not detected or found to be very low (Fig. 3). This expression pattern is essentially the same as that shown for *Arabidopsis PI* gene (Goto and Meyerowitz, 1993). Genomic sequences of *MdPI* were amplified using the PCR primers P5 within the MADS-box and P6 within the 3' non-translated region. Two clones containing the *MdPI* genomic DNA were sequenced and found to contain the same sequences having six easily identifiable introns. The relative positions of intron 2 to intron 6 are highly conserved compared to the positions of 5 introns in *PI* gene (Fig. 2). We conclude that *MdPI* is the *PI* homolog based on these results having highest sequence identity and conserved intron positions and mRNA expression patterns.

In an experiment to examine whether there is a mutation in *MdPI* of Rae Ime, the expression level of *MdPI* in flower buds was determined. Expression of *MdPI* in the apetalous Rae Ime flower buds is not detected, but is readily detected in normal flower buds of the Granny Smith variety (Fig. 3). In *Arabidopsis pi* mutants, *PI* expression is reduced or abolished in flower buds (Goto and Meyerowitz, 1994).

A second experiment compared RFLP patterns for Rae Ime with normal apple cultivars using the *MdPI* cDNA as a probe. Southern hybridisation shows different RFLP patterns between Rae Ime and Granny Smith with both *Eco*RI and *Hind*III digestion (Fig. 4) although Granny Smith RFLP pattern is conserved in another apple variety Royal Gala (data not shown). Both the expression and RFLP data indicate that the *MdPI* gene in Rae Ime has been mutated. As both enzyme digestions reveal RFLP differences, the mutation is likely to be a gross change in gene structure rather than a point mutation

Genomic DNA fragments were cloned from Granny Smith and Rae Ime using two primers P3 and P7 designed with *MdPI* cDNA sequence. The Rae Ime fragments were 11 kb while the Granny Smith fragments were 2 kb (Fig. 5a). These fragments show a hybridisation signal to the *MdPI* cDNA probe (Fig. 5b). Clones containing these fragments were partially sequenced from two ends. The Rae Ime fragments have the same sequence to the Granny Smith fragments at two ends, but with an insertion in the intron 4 of *MdPI* gene in Rae Ime (Fig. 5b). The insertion sequences showed no homology to DNA sequences in GeneBank. This result confirmed that there is a mutation in the *MdPI* gene in Rae Ime.

By way of confirmation that it is the mutation of the *MdPI* gene which is responsible for the parthenocarpic phenotype, the *MdPI* gene from two further parthenocarpic apple varieties, Spencer Seedless and Wellington Bloomless, was sequenced (data not shown). This revealed an approximately 9 kb insertion in each gene. Thus, in the three parthenocarpic apple varieties examined, there are two different insertion sites in the *MdPI* gene both of which lead to the parthenocarpic phenotype. Spencer Seedless and Wellington Bloomless have the same insertion site, which is different from that in Rae Ime. These confirmatory results demonstrate that independent mutations in *MdPI* generate the same apetalous/parthenocarpic phenotype.

The difference in fruit development between Rae Ime apple and *pi Arabidopsis* may be explained in two different ways. Firstly, *MdPI* may have different function compared to *PI* in influencing ovary and fruit development. Sufficient functional differences have been shown for homologs of floral homeotic genes in different plant species (Causier *et al.*, 1999). Secondly, apple fruit develops from both ovary and the fused bases of sepals, petals and stamens (Pratt, 1988). Apple differs from tomato and *Arabidopsis*, two model systems often used in studies of fruit development, where the fruit or silique develops from ovary tissue only (Weigel and Mererowitz, 1994; Gillaspy *et al.*, 1993). The differences in fruit structure may cause different fruit development after a mutation in a floral homeotic gene.

INDUSTRIAL APPLICATION

In its primary aspect, the invention has application in modulating, and in particular reducing or eliminating seed-bearing capacity in fruiting plants. Such plants have utility in horticulture.

It will also be possible to employ the polynucleotides of the invention in breeding programmes to monitor the progress made towards breeding a stable seedless fruiting plant.

The availability of reproductively null or sterile trees has the additional advantage that it will be possible to introduce further exogenous genetic material into those trees without the risk that the material will be passed on to other trees.

Those persons skilled in the art will appreciate that the specific description provided is exemplary only, and that modifications and variations may be made without departing from the scope of the invention.

REFERENCES

Brase, K.D. The vascular anatomy of the flower of Malus domestica Borkh. f. apetala Van Eseltine. M. S. Thesis. Cornell University (1937).

Causier, B., Weir, I. & Davies B. MADS-box factors in hermaphrodite flower development. In: Ainsworth, C.C. (ed) Sex Determination in Plants, BIOS Scientific Publishers Ltd, Oxford. pp1-23 (1999).

Chan, B.G. & Cain, J.G. The effect of seed formation on subsequent flowering in apple. Proc. Amer. Soc. Hort. Sci. 91: 63-68 (1967).

Chang, S., Puryear, J. & Cairney, J. A simple and efficient method for isolating RNA from pine trees. Plant Mol. Biol. Rep. 11:113-116 (1993).

Dong, Y-H, Janssen, B.-J., Bieleski, L.E.F., Atkinson, R.G., Morris, BA M. & Gardner, R.G. Isolating and characterizing genes differentially expressed early in apple fruit development. J. Amer. Soc. Hort. Sci. 122:752-757 (1997).

Gillaspy, G., Ben-David, H. & Gruissem, W. Fruits: a developmental perspective. Plant Cell 5:1439-1451 (1993).

Goonewardene, H.F., Kwolek, W.F. & Hayden, R.A. Survival of immature stages of the codling moth (Lepidoptera: Tortricidae) on seeded and seedless apple fruit. J. Econ. Entomol. 77:1427-1431 (1984).

Goto, K. & Meyerowitz E.M. Function and regulation of the Arabidopsis floral homeotic gene PISTILLATA. Genes Dev 8:1548-1560 (1994).

Jack, T., Brockman, L.L. & Meyerowitz, E.M. The homeotic gene APETALA3 of Arabidopsis thaliana encodes a MADS box and is expressed in petals and stamens. Cell 68:683-697 (1992).

Marshall and Hodgson (1998) *Nature Biotechnology* 16:27-31.

Pratt, C. Apple flower and fruit: morphology and anatomy. Hort. Rev. 10:273-307 (1988).

Rogers, S.O & Bendich, A.J. Extraction of DNA from plant tissue. In: Gelvin, S.B & Schilperoort, R.A. (eds) Plant Molecular Biology Manual. Kluwer Academic Publishers, Dordrecht, Belgium, pp. A6:1-13 (1988)

Tobutt, K.R. Combining apetalous parthenocarpic with columnar growth habit in apple. Euphytica 77:51-54 (1994).

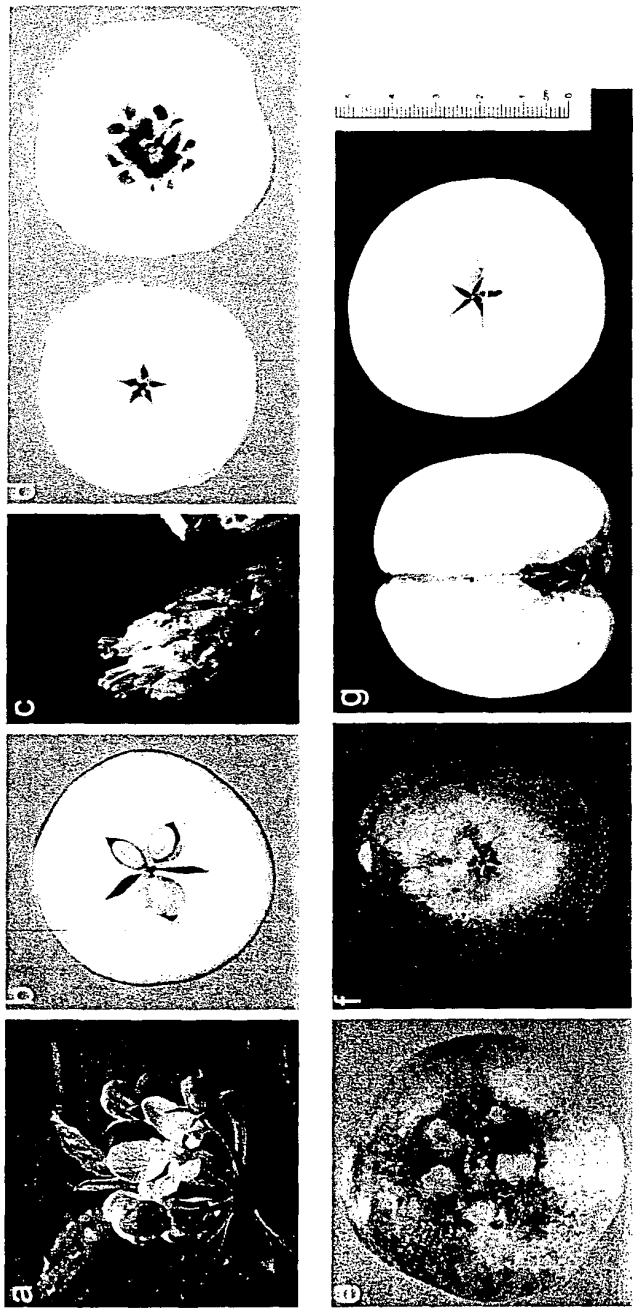
Weigel, D. & Meyerowitz, M. The ABCs of floral homeotic genes. Cell 78:203-209 (1994).

WEST WALKER BENNETT

per

ATTORNEYS FOR THE APPLICANT

Fig. 1



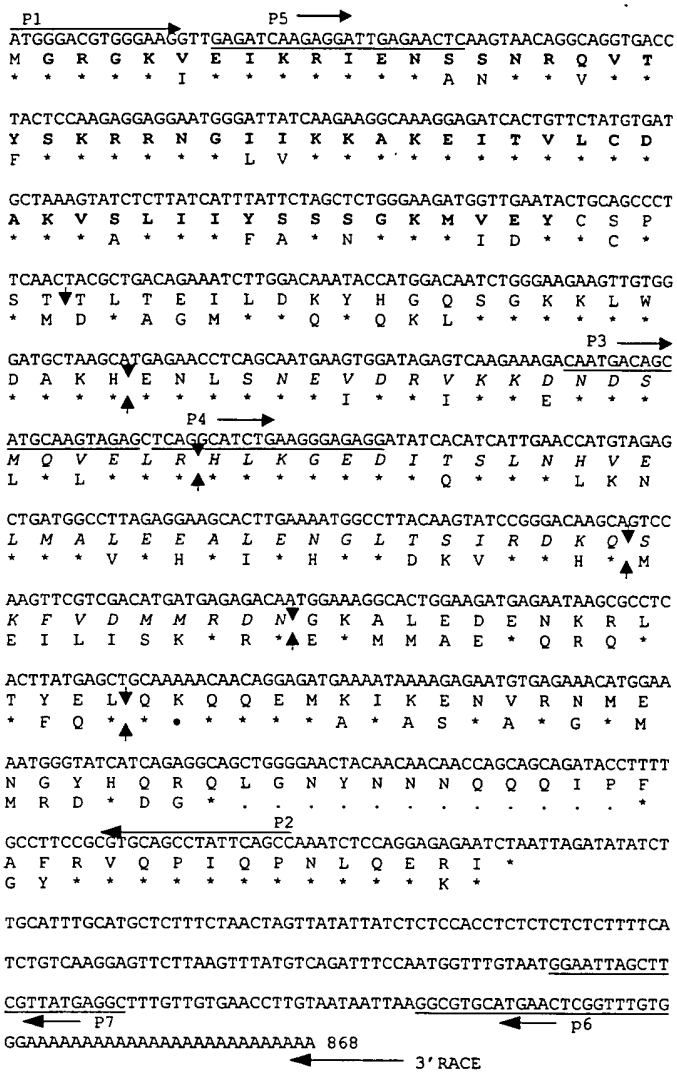


Fig. 2

Fig. 3

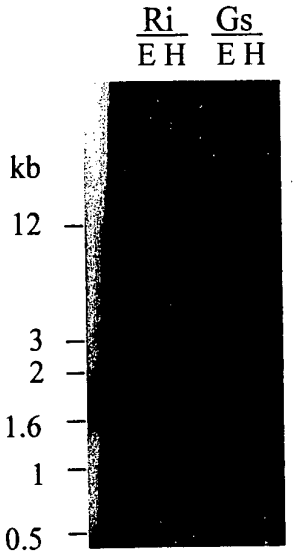
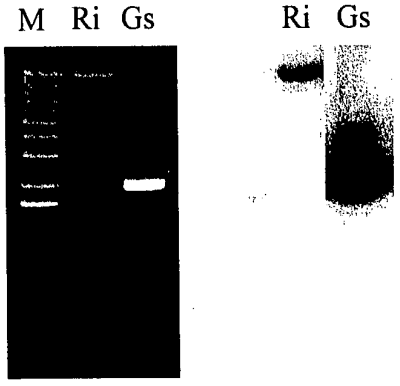


Fig. 4




```
Ri5'  ATATATGTTTATAGTTATAT TGTCACAGCCCGTTCCAAAT
Gs    ATATATGTTTATAGTTATAT GTTGTTGTATACTTAATTGT
Ri3'  GGGCGGGGTGTGTCATATAT GTTGTTGTATACTTAATTGT
```

c

Fig. 5